

GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 Compugen Ltd.  
**OM protein - protein search, using sw model**  
 Run on: June 3, 2003, 15:04:33 ; Search time 18 Seconds  
 (without alignments)  
 1276.451 Million cell updates/sec

Title: US-09-887-784-4  
 Perfect score: 1274  
 Sequence: 1 MVSKEELFTGVVPLIVLD.....VLIGFVTAAGITLGMDELYK 239  
 Scoring table: BL05N62  
 Gapop 10.0 , Gapext 0.5  
 Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing first 45 summaries

Database : PIR73:  
 1: Pir1:  
 2: Pir2:  
 3: Pir3:  
 4: pir4:  
 A: Residues: 1-107, 'S', 109-238 <PRA1>  
 A: Cross-references: GB:MG2654; NID:9155662; PIDN:AAA27722.1; PID:9155663

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1241	97.4	238	1	JO1514		green fluorescent hypothetical prote
2	105	8.2	78	2	H72228		leucine-tRNA ligas
3	91.5	7.2	881	2	H64102		DNA-binding protein
4	90	7.1	632	2	T06586		leucyl-tRNA synthetase
5	89.5	7.0	887	2	E82590		DNA topoisomerase
6	88.5	6.9	655	2	D83917		iron-sulfur cofact
7	87.5	6.9	370	2	E70390		leucyl-tRNA synthetase
8	87.5	6.9	860	2	AC0582		DNA-directed DNA polymerase
9	87.5	6.9	2222	1	A36028		hypothetical prote
10	87.5	6.9	2573	2	D71614		dihydrofolamide d
11	87	6.8	578	1	I40794		S-layer protein protective surface
12	87	6.8	874	2	JC4930		protective surface
13	86.5	6.8	797	2	JCA078		synaptogamin O-p55
14	86.5	6.8	808	2	F64102		DNA-directed DNA polymerase
15	86	6.8	357	2	G81355		inter-alpha-trypsin
16	85.5	6.7	425	2	C97354		cellulase (EC 3.2.1.2)
17	85.5	6.7	788	1	JDNTHH		hypothetical prote
18	85.5	6.7	889	2	JC5576		oligoendopeptidase
19	85	6.7	941	2	S29043		imidazoleglycerol-
20	85	6.7	281	2	AD2052		DNA-directed DNA polymerase
21	84.5	6.6	439	2	JHO414		IgA FC receptor
22	84.5	6.6	613	2	A99552		IgA FC receptor
23	84	6.6	353	2	EB8941		photosensory protein
24	83.5	6.6	836	1	JDVLD		hypothetical prote
25	83.5	6.6	1134	2	A6034		alpha-amylase homo
26	83.5	6.6	1164	1	FC504G		alpha-amylase homo
27	83	6.5	461	2	T27556		alpha-amylase homo
28	83	6.5	471	2	T06532		alpha-amylase homo
29	83	6.5	774	2	T39533		alpha-amylase homo

#### ALIGNMENTS

#### RESULT 1

JO1514 green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)  
 C; Species: Aequorea victoria  
 C; Date: 03-Dec-1999 \*sequence\_revision 03-Dec-1999 #text\_change 23-Mar-2001  
 C; Accession: JS0692; JO1514; PG0335; S4893; S51330; S51331  
 R; Prasher, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J.  
 Gene ID: 229-233; 1992  
 A; Title: Primary structure of the Aequorea victoria green-fluorescent protein.  
 A; Reference number: JO1514; MUID:92175527; PMID:1347277  
 A; Accession: JS0692  
 A; Molecule type: DNA  
 A; Residues: 1-107, 'S', 109-238 <PRA1>  
 A; Cross-references: GB:MG2654; NID:9155662; PIDN:AAA27722.1; PID:9155663  
 A; Accession: Q1514  
 A; Molecule type: mRNA  
 A; Residues: 1-99, 'F', 101-140, 'L', 142-218, 'V', 220-238 <PRA2>  
 A; Cross-references: GB:MG2653; NID:9155660; PIDN:AAA27721.1; PID:9155661  
 A; Accession: P0335  
 A; Molecule type: Protein  
 A; Residues: 46-64; 74-122; 132-151; 154-183; 185-200 <PRA3>  
 R; Inouye, S.; Tsujii, P.T.  
 FBB5 Lett. 351, 211-214, 1994  
 A; Title: Evidence for redox forms of the Aequorea green fluorescent protein.  
 A; Reference number: S48693; MUID:94364470; PMID:8032767  
 A; Accession: S48693  
 A; Molecule type: mRNA  
 A; Residues: 1-24, 'O', 26-156, 'P', 158-171, 'K', 173-238 <IN0>  
 A; Cross-references: GB:L29345; NID:9606383; PMID:AAA58246.1; PID:9606384  
 R; Watkins, J.N.; Campbell, A.K.  
 A; Molecule type: mRNA  
 A; Residues: 1-13, 'V', 15-24, 'Q', 26-44, 'N', 46-153, 'G', 155-156, 'P', 158-171, 'K', 173-227  
 A; Cross-references: ENBL:X39595; NID:9634008; PMID:AAA58189.1; PID:9634009  
 A; Accession: S51330  
 A; Reference source: clone gfp1  
 A; Accession: S51331  
 A; Molecule type: mRNA  
 A; Residues: 1-24, 'Q', 26-29, 'R', 31-83, 'L', '85-153, 'G', 155-156, 'P', 158-171, 'K', 173-208  
 A; Cross-references: ENBL:X395960; NID:9634010; PMID:9634011  
 A; Reference number: A65692; PDB:1GF1  
 A; Content: annotation, X-ray crystallography, 1.9 angstroms, residues 'A', 279, 'R'  
 A; Note: engineered sequence based on JQ1514, cloned and expressed in Escherichia coli  
 R; Yang, F.; Moss, L.G.; Phillips Jr., G.N.  
 Natl. Biotechnol. 14, 1246-1251, 1996  
 A; Title: The molecular structure of green fluorescent protein.  
 A; Reference number: A58953; MUID:98294543; PMID:9631087





A; Reference number: S60989  
 A; Accession: SGD919  
 A; Molecule type: DNA  
 A; Residues: 1-2221 <SEND>  
 A; Cross-references: EMBL:X92494; NID:91045236; PID:91045247  
 R; Sen-Gupta, M.; Lyck, R.; Niedenthal, R.K.; Fleig, U.N.; Hegemann, J.H.  
 submitted to the Protein Sequence Database, April 1996  
 A; Reference number: S63235  
 A; Accession: S63235  
 A; Molecule type: DNA  
 A; Residues: 1-2222 <SEND>  
 A; Cross-references: EMBL:Z71538; NID:91302316; PID:91302317; GSDB  
 A; Experimental source: strain S28BC  
 R; Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.K.; Hegemann, J.H.  
 yeast 12, 505-514, 1996  
 A; Title: The sequence of a 24152 bp segment from the left arm of chromosome XIV from  
 A; Reference number: S65111; MUID:96310631; PMID:8740425  
 A; Accession: S65111  
 A; Status: nucleic acid sequence not shown; translation not shown  
 A; Molecule type: DNA  
 A; Residues: 1-2221 <SEND>  
 A; Cross-references: EMBL:X92494; NID:91045236; PID:91045247  
 A; Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995  
 C; Genetics  
 A; Gene: SGD:POL2; DUN2; MIPS:YNL252w  
 A; Cross-references: SGD:S0005206; MIPS:YNL262w  
 A; Map position: 14L  
 C; Superfamily: DNA-directed DNA polymerase II  
 C; Keywords: DNA binding; nucleotidyltransferase; nucleus; zinc finger  
 Query Match Score 87.5; DB 1; Length 2222;  
 Best Local Similarity 28.2%; Pred. No. 1e+02; Gaps 31; Gaps 7;  
 Matches 37; Conservative 14; Mismatches 49; Indels 19  
 Qy 54 LPVPPNP-TLVTTPLSIVQVOCFSRYPDPM-----KQHDEFFKSAMPPEGYY---QERTI 99  
 Db 883 LPKSFPTPYTFLENRKLYLSPYKSLNRYHQKEFTNHQOELPKNLYVYETHSNT 942  
 Qy 100 FFKDDGNYKTR--AEVKEFGDTLVNR-----IELKGIDFEDGNLGHKLEYNN 147  
 Db 943 FFEVDGPYKAMILPPSSKEEGRGKIKRYAVFNEDGSLAEKGPFELKRQGEL--QLIKNEQ 999  
 Qy 148 S--HNVYIMAD 156  
 Db 1000 SDIFKVFLGED 1010  
 RESULT 10  
 D71614 hypothetical protein PFB0460c - malaria parasite (Plasmodium falciparum)  
 C; Species: Plasmodium falciparum  
 C; Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 21-Jul-2000  
 C; Accession: D71614  
 R; Gardner, M.J.; Tettelin, H.; Carruccio, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.; Partea, M.; Salzberg, S.; Zhou, L.; Sutton, G.C.; Clayton, R.; White, O.; Smith, Science 282, 1126-1132, 1998  
 A; Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum  
 A; Reference number: A71600; MUID:93021743; PMID:9844551  
 A; Accession: D71614  
 A; Status: Preliminary; nucleic acid sequence not shown; translation not shown  
 A; Molecule type: DNA  
 A; Residues: 1-573 <GAR>  
 A; Cross-references: GB:AB001396; GB:AE001362; NID:93845188; PIDN: AAC71881.1; PID:93  
 A; Experimental source: clone 3D7  
 C; Genetics:  
 A; Gene: PFB0460c  
 Query Match Score 87.5; DB 2; Length 2573;  
 Best Local Similarity 26.2%; pred. No. 1.2e+02;  
 Matches 34; Conservative 30; Mismatches 53; Indels 13; Gaps 5;  
 Qy 94 VQERTIFFKD--DGNYKTRAEVKFEDDTLVRIELGIDFEDGNLGHKLEYNN-YNSH 149  
 C; Genetics:  
 A; Gene: B36028  
 A; Molecule type: protein  
 A; Residues: 1214-1216, 'X', 1218-1221 <D02>  
 R; Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.N.; Hegemann, J.M.  
 submitted to the EMBL Data Library, October 1995  
 A; Title: The sequence of a 24152 bp segment from the left arm of chromosome XIV from  
 A; Reference number: A36028; MUID:90381771; PMID:2169349  
 A; Accession: A36028  
 A; Molecule type: DNA  
 A; Residues: 1-2222 <MOR>  
 C; Genetics:  
 A; Cross-references: B36028  
 A; Molecule type: protein  
 A; Residues: 1214-1216, 'X', 1218-1221 <D02>  
 R; Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.N.; Hegemann, J.H.  
 submitted to the EMBL Data Library, October 1995  
 A; Title: The sequence of a 24152 bp segment from the left arm of chromosome XIV from  
 A; Reference number: A36028; MUID:90381771; PMID:2169349  
 A; Accession: A36028  
 A; Molecule type: DNA  
 A; Residues: 1-2222 <MOR>

Db	126 LKKETILCKDIKSGSNNDPMDEISLFLKDDMVDDEKLK - DPEKSSLKIKNKEVYNNFIYKMK 183	A; Residues: 1-874 <ZHU> A; Cross-references: GB:U38842; NID:91055336; PIDN: AAC44405.1; PMID:91055337 A; Experimental source: strain NM105 C; Comment: This protein is a glycoprotein. It functions as protective coats, molecule C; Genetics: A; Gene: opa C; Superfamily: S-layer repeat homology F; 1/29/Domain: signal sequence #status predicted <SIG> F; 30-874/Product: S-layer protein #status predicted <MAT> F; 157-209/Domain: S-layer repeat homology <SLR3>
Qy	150 NYKIMADKOKNGKIVKFIRHNTEDGSKYOLADHYQONTPIGDSPVLLPDNHYLSTOSALS 209	
Db	184 NLHKENKKDKKEKKNKHNNDENNRM---IYKRN1---DKTHYILDNNVYHILNDIN 236	
Qy	210 KDPNEKRDHM 219	
Db	237 TYLKRDHM 246	
<b>RESULT 11</b>		
	I40794 dihydrolipoamide dehydrogenase (EC 1.8.1.4) [validated] - Clostridium magnum dihydrolipoamide dehydrogenase complex E3; acetoin dehydrogenase	
N; Alternative names: 2-oxoglutarate dehydrogenase complex chain E3; S-complex 50K chain	hydrogenase	
C; Species: Clostridium magnum	hydrogenase	
C; Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 *text_change 21-Jul-2000		
C; Accession: I40794		
R; Kruger, N.; Oppermann, F.-B.; Lorenzl, H.; Steinbuchel, A.		
A; Title: Biochemical and molecular characterization of the Clostridium magnum acetoin dehydrogenase		
A; Reference number: I40789; PMID:94266715; PMID:8206840		
A; Accession: I40794		
A; Status: preliminary; translated from GB/EMBL/DDBJ		
A; Molecule type: DNA		
A; Residues: 1-578 <KR>		
A; Cross references: GB:L31844; NID:9472324; PIDN:AAA21748.1; PID:9472330		
C; Function:		
A; Description: catalyzes the oxidation of dihydrolipoamide to Lipoinamide using NAD		
A; Pathway: acetoin dehydrogenase enzyme system		
A; Superfamily: Alloalloxazine dihydrolipoamide dehydrogenase; dihydrolipoamide dehydrogenase		
C; Keywords: FAD; flavoprotein; Lipoinamide; NAD; oxidoreductase; redox-active disulfide		
F; 117-155/Region: biotin-binding homology <RPB>		
F; 119-561/Region: beta alpha-beta FAD nucleotide-binding fold		
F; 287-315/Region: dihydrolipoamide dehydrogenase homology <LDL>		
F; 153-158/Disulfide bonds: beta alpha-beta NAD nucleotide-binding fold		
<b>RESULT 13</b>		
	JC4078 protective surface antigen D-15 Precursor - Haemophilus influenzae (type b)	
Query Match Score 87; DB 1; Length 578;		
Best Local Similarity 23.5%; Pred. No. 19; Mismatches 40; Indels 52; Gaps 12;		
Matches 54; Conservative 40; Mismatches 84; Indels 52; Gaps 12;		
Qy	10 TGVYPLVLEDGDNHGKPSVSGEGDATGKYLTK-----FICTGKLPVPMPVLT 63	A; Variety: type b C; Species: Haemophilus influenzae C; Date: 30-Jun-1995 *sequence_revision 14-Jul-1995 *text_change 29-Sep-1999
Db	255 TGSMPFPIPE---GNKLS---CVIDST-GAISLESNPESIALIGGGTGVFAFISFN 305	C; Accession: JC1078 R; Flack, F.S.; Loosmore, S.; Chong, P.; Thomas, W.R. Gene 156, 97-99, 1995 A; Title: The sequencing of the 80-kDa D15 protective surface antigen of Haemophilus
Qy	64 TLSKGVQCFSRYPDHMKQHDFFKSAMPEGVQERTIFFKDGNYKTRAEVKFEGDTLYNR 123	A; Reference number: JC4078; PMID:95255676; PMID:737523
Db	306 SLGCKVSTIEMLPHLPPMDREISI-----AKAKLIRDGININNN 346	A; Molecule type: DNA A; Residues: 1-797 <FLA> A; Cross references: GB:U13961; NID:9537447; PIDN: AAA85645.1; PMID:9537448 A; Experimental source: type b C; Superfamily: protective surface antigen D-15
Qy	124 IELKGIDFKEDG---NLIGKRLYENSHAVYIMADKQKN - GIVRNFKDRHNIEGSGVQ 178	C; Superfamily: protective surface antigen D-15 C; Keywords: surface antigen
Db	347 CKVTRLEQGEDGLKYSFIGDKGEESIDVERVLYIAGRRSNTIEGLDVE-KIGVKTBGSII 405	F; 1-19/Domain: signal sequence #status predicted <SIG> F; 20-79/Product: protective surface antigen D-15 #status predicted <MAT>
Qy	179 LADHYQONT---PIGD---GPVLDPDNHYLSTOSALSKD---PNEKRDH 218	
Db	406 VNDRMEINVEISIAGDCGTGKMLA - HVRSQGTVAARENINGQNKKMDY 453	Qy Match Score 86.5%; Pred. No. 32; Mismatches 78; Indels 63; Gaps 11;
<b>RESULT 12</b>		
	JC4910 S-layer protein precursor - Bacillus licheniformis	
C; Species: Bacillus licheniformis		
C; Date: 22-Oct-1996 #sequence_revision 01-Nov-1996 *text_change 15-Oct-1999		
C; Accession: JC4930		
R; Zhu, X.; McVeigh, R.R.; Malathi, P.; Ghosh, B.K.		
Gene 173, 189-194, 1996		
A; Title: The complete nucleotide sequence of the Bacillus licheniformis NM105 S-layer protein		
A; Reference number: JC4930; PMID:97082965; PMID:8964497		
A; Status: preliminary		
A; Molecule type: DNA		

Qy 197 P--DNHILSTQSALSKDPEKRDHMVLLGFVTAAGITLG 233  
 Db 594 PGSDNKKYKLSADQGFPLDRDHMWWVSAKASAGTANG 632

**RESULT 14**  
**F64102** protective surface antigen D-15 - Haemophilus influenzae (strain Rd KW20)  
 C;Species: Haemophilus influenzae  
 C;Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 18-Sep-1998  
 R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage,  
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.T.; Glodek, A.; Kelley, J.M.; Weidman,  
 , D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, N.S.M.  
 Science 269, 496-512, 1995  
 A;Authors: Gnealm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Wenter,  
 A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A;Reference number: A64000; PMID:7542800  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-808 <TIGR>  
 A;Cross-references: GB:L42023; TIGR:HI0917  
 C;Superfamily: protective surface antigen D-15  
 C;Keywords: surface antigen

Query Match 6.8%; Score 86.5; DB 2; Length 808;  
 Best Local Similarity 21.9%; Pred. No. 32; Mismatches 30; Indels 63; Gaps 11;  
 Matches 48; Conservative 30; Mismatches 78; Indels 63; Gaps 11;

Qy 65 LSYGVQCFSRYPDHMKOHDF-----FVKSAMPEGIVQE-----RTI 99  
 Db 440 IGYGTTSGISQASVQDNEFLGTGAASIAGTKNDGTSVNGLGXTEPYFTKDGVSLLGVN 499

Qy 100 FFKDDGNYKTRAEVKEFGDTLVNRIEJRGIDFKEDGNT---LGH-----KLEYNNNS 148  
 Db 500 FFENTDNSKSDTSSNKRTTGSNTL-GFPYNENNNSYYVGLHTYNIKTSNFALIYN--- 555

Qy 149 HNVYIMADKQR-NGIKVNFKIRHNIEEDGSVOLADHQQ-----NTPIGDGPYLL 196  
 Db 556 RNLYIQSMKFFGNGIIN-----DFDSFGWNTAISLNRGXFPTKGVKASLG-GRVTI 606

Qy 197 P--DNHILSTQSALSKDPEKRDHMVLLGFVTAAGITLG 233  
 Db 607 PGSDNKKYKLSADQGFPLDRDHMWWVSAKASAGTANG 645

**RESULT 15**

G81355 tRNA (uracil-5')-methyltransferase (EC 2.1.1.35) C]0831c [Imported] - Campylobacter jejuni  
 C;Species: Campylobacter jejuni  
 C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
 C;Accession: G81355  
 R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Challacombe,  
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; Vanlylet, A.; Whitehead, S.; Barrell,  
 B.; Nature 403, 665-668, 2000  
 A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
 A;Reference number: A81250; PMID:10150912; PMID:10688204  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-357 <PAR>  
 A;Cross references: GB:AL1139076; GB:AL111168; NID:96968128; PID:CAB73096.1; PID:9696827  
 A;Experimental source: serotype O2, strain NCTC 11168  
 C;Genetics:  
 A;Gene: trmA; C]0831c  
 C;Keywords: methyltransferase; S-adenosylmethionine

Query Match 6.8%; Score 86; DB 2; Length 357;  
 Best Local Similarity 24.8%; Pred. No. 12; Mismatches 18; Indels 34; Gaps 5;  
 Matches 30; Conservative 18; Mismatches 39; Indels 34; Gaps 5;  
 C;Keywords: methyltransferase; S-adenosylmethionine

Qy 80 KQHDFEKSAMPEGIVQERTIFFKDDGNYKTRAEVKF--EGDTLV-----NRIELKG 128